Abstract Title Page
Not included in page count.

Title: "Theoretical and Empirical Underpinnings of the What Works Clearinghouse Attrition Standard for Randomized Controlled Trials"

Authors and Affiliations:

John Deke, Mathematica Policy Research (led the development of the WWC attrition standard)

Hanley Chiang, Mathematica Policy Research

Abstract Body

Limit 4 pages single-spaced.

Background / Context:

Description of prior research and its intellectual context.

Sample attrition in randomized controlled trials (RCTs) can lead to biased impact estimates. This is because the process by which individuals leave the study sample could be related to both treatment status and the outcomes of interest, leading to systematic differences between the individuals remaining in the treatment and control groups. While attrition *might* occur in a way that is unrelated to outcomes and treatment status, it is rarely possible to *know* this since the researcher typically does not control or understand the process by which individuals leave the study sample. Consequently, findings from RCTs that experience sample attrition should be regarded as less credible than findings from RCTs that experience no sample attrition. The fact that missing outcomes data can lead to bias in general (not just in RCTs) is a well-known issue that has received considerable attention in the methodological literature (Rubin 1976; Little and Rubin 2002; Puma et al 2009).

To limit the potential for attrition bias to influence assessments of intervention effectiveness, several federally sponsored systematic evidence reviews have established a standard for acceptable levels of sample attrition in RCTs. These evidence reviews include the What Works Clearinghouse (WWC), the Home Visiting Evidence of Effectiveness Review (HomVEE), and the Teen Pregnancy Prevention Evidence Review (TPPER) [see http://ies.ed.gov/ncee/wwc/, http://homvee.acf.hhs.gov/, and http://www.hhs.gov/ash/oah/oah-initiatives/teen_pregnancy/db/].

The attrition standard developed for the WWC was the progenitor of the other two standards. Initially the WWC attrition standard consisted of arbitrary cutoffs with respect to the overall attrition rate and the difference in attrition rates between the treatment and control groups. This initial standard had no empirical basis and no theoretical basis beyond the basic understanding that less attrition is likely preferable to more attrition. However, starting with version 2.0 of the WWC Procedures and Standards Handbook (http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=19), the WWC attrition standard moved from being a set of arbitrary cutoffs to being a range of cutoffs based on a mathematical model of attrition bias with the choice of key parameters informed by empirical data analysis.

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

The purpose of this paper/presentation is to explain the WWC attrition model, the process for selecting key parameter values for that model, and how the model informed the development of the WWC attrition standard and can be used to develop attrition standards tailored to other substantive areas.

Setting:

Description of the research location.

(May not be applicable for Methods submissions)

Not applicable

Population / Participants / Subjects:
Description of the participants in the study: who, how many, key features, or characteristics.

(May not be applicable for Methods submissions)

Not applicable

Intervention / Program / Practice:
Description of the intervention, program, or practice, including details of administration and duration.

(May not be applicable for Methods submissions)
Not applicable

Significance / Novelty of study:

Description of what is missing in previous work and the contribution the study makes.

While others have developed approaches to bounding bias in a variety of contexts (for example, Manski 1990), we believe the WWC attrition standard may constitute the first use of model-based, empirically-supported bounds on attrition bias in the context of a federally-sponsored systematic evidence review. Previously federal agencies such as OMB and NCES have used standards for response rates in data collection efforts, but these standards are not based on estimates of expected bias, but rather on what response rates have been feasible to achieve in the past (http://www.whitehouse.gov/omb/inforeg/pmc_survey_guidance_2006.pdf and http://nces.ed.gov/pubs2003/2003601.pdf).

Meeting the WWC attrition standard (or one of the attrition standards based on the WWC standard) is now an important consideration for researchers conducting studies that could potentially be reviewed by the WWC (or other evidence reviews).

Understanding the basis of this standard is valuable for anyone seeking to meet existing standards and for anyone interested in adopting this approach to developing a standard (that is, combining a theoretical model with empirical estimates of key parameters) in a new context.

Statistical, Measurement, or Econometric Model:

Description of the proposed new methods or novel applications of existing methods.

When attrition is related to outcomes, different rates of attrition between the treatment and control groups can lead to biased impact estimates. Furthermore, if the relationship between attrition and outcomes differs between the treatment and control groups, then attrition can lead to bias even if the attrition rate is the same in both groups. The focus here is to specify a model showing how bias depends on the correlation between outcomes and attrition and the combination of overall and differential attrition in an RCT.

To set up the model, consider a variable representing an individual's latent (unobserved) propensity to respond, z. Assume z is normally distributed with mean zero and standard deviation

one. If the proportion of individuals who respond is P, an individual is a respondent if his or her value of z exceeds a threshold, z^*:

$$(1) \quad z > \Phi^{-1}(1-P) \equiv z^*$$

where Φ is the standard normal cumulative distribution function. For example, in a scenario in which 75 percent of individuals respond ($P = 0.75$), an individual is a respondent if his or her value of z exceeds the value corresponding to the 25th percentile in the z distribution [that is, exceeds $\Phi^{-1}(1-0.75)$].

The outcome at follow-up, y, is the key variable of interest. We assume that y has a normal distribution. Moreover, we assume that y has mean zero and standard deviation one, given that any variable can be standardized in this way. The relationship between y and z can then be modeled as

$$(2) \quad y = \alpha * z + u$$

where α is the correlation between z and y, and u is a random variable that is independent of z.[1] Note that there are no covariates and the model assumes no effect of the treatment on the mean outcome. If α is 1 or -1, the entire outcome is explained by the propensity to respond. If α is zero, none of the outcome is explained by the propensity to respond, which is the case when attrition is completely random.

The correlation between the propensity to respond and outcomes may differ by treatment status. Therefore, we specify Equation (2) separately for treatment and control group members (subscripted by t and c, respectively):

$$(3) \quad \begin{aligned} y_t &= \alpha_t * z_t + u_t \\ y_c &= \alpha_c * z_c + u_c . \end{aligned}$$

Because there is no true impact of the intervention on mean outcomes in this model, an unbiased estimator of the impact should, in expectation, find no difference in outcomes between the treatment and control groups. Therefore, in the presence of attrition, *bias* is equal to the difference between the expected values of y_t and y_c among respondents. Based on the properties of truncated normal distributions, the analytic formula for the bias, B, is

$$(4) \quad \begin{aligned} B &= E(y_t \mid z_t > z_t^*) - E(y_c \mid z_c > z_c^*) \\ &= \alpha_t E(z_t \mid z_t > z_t^*) - \alpha_c E(z_c \mid z_c > z_c^*) \\ &= \frac{\alpha_t \times \phi(\Phi^{-1}(1-P_t))}{P_t} - \frac{\alpha_c \times \phi(\Phi^{-1}(1-P_c))}{P_c} \end{aligned}$$

[1] In order for y to be a N(0,1) variable, u must be normally distributed with mean zero and standard deviation $\sqrt{1-\alpha^2}$.

where ϕ is the standard normal density function.

Equation (4) shows that bias can be generated by treatment-control differences in the response rates (P_t and P_c) or in the correlation between y and z (α_t and α_c). If neither P nor α differs between these groups, then there is no bias because the same kinds of individuals respond from both groups.[2] However, if response rates differ between the treatment and control groups ($P_t \neq P_c$), then bias occurs even when $\alpha_t = \alpha_c$, because respondents in the treatment and control groups have different average values of z and, thus, different average values of y. Moreover, if $\alpha_t \neq \alpha_c$, then impact estimates will be biased even if the response rate is the same in both groups; respondents from the two groups will have different average values of y stemming from the differences in α.[3]

Usefulness / Applicability of Method:

Demonstration of the usefulness of the proposed methods using hypothetical or real data.

The model specified above can be used to calculate the expected bias for any combination of overall and differential given estimates of α_t and α_c. With those estimates of bias, we can then construct bounds on acceptable rates of overall and differential attrition. Here we (1) describe the process for estimating α_t and α_c and (2) illustrate attrition bounds corresponding to particular values of α_t and α_c.

1. Estimating α_t and α_c

If we were able to observe outcomes for both respondents and non-respondents in the treatment and control groups we could calculate α_t and α_c directly. Let Δ_g denote the difference in outcomes, in effect size units, between respondents and nonrespondents in group g (either the treatment [t] or control [c] group). It can be shown that

$$(5a)\ \ \Delta_g = E(y_g \mid z_g > z_g^*) - E(y_g \mid z_g \leq z_g^*) = \frac{\alpha_g \times \phi(\Phi^{-1}(1-P_g))}{P_g(1-P_g)},$$

which implies that

$$(5b)\ \ \alpha_g = \frac{\Delta_g P_g(1-P_g)}{\phi(\Phi^{-1}(1-P_g))}.$$

[2] Those who attrite, nonetheless, will differ systematically from those who do not attrite, which may compromise the external validity of the study. However, we do not address that issue here.

[3] It is possible that a difference in the rate of attrition between groups could offset a difference between α_t and α_c. However we conservatively assume the opposite—that these differences are reinforcing, not offsetting.

Of course, we cannot observe outcomes for nonrespondents, so we cannot observe Δ_g directly. However, in studies that have both follow-up and baseline test scores, we can use the baseline test scores as proxies for the follow-up test scores because baseline scores are typically correlated with follow-up scores. Therefore we can use the difference in *baseline* test scores between respondents and nonrespondents as the proxy for Δ_g, and we use Equation (5b) to calculate α_g for $g = t, c$.

2. Bounding Attrition Bias

Given estimates of α_t and α_c, we can calculate expected bias for every combination of overall and differential attrition rates. We can then assess whether the expected bias for each combination of overall and differential attrition rates exceeds a maximum tolerable level (for the WWC, the level was set at 0.05 standard deviations). Figure 1 illustrates the combinations of overall and differential attrition that yield expected bias below the maximum tolerable amount for two different sets of assumptions for α_t and α_c. The green/bottom-left region shows combinations of overall and differential attrition that yield attrition bias less than 0.05 under pessimistic assumptions ($\alpha_t = 0.45$ and $\alpha_c = 0.39$), the yellow/middle region shows additional combinations that yield attrition bias less than 0.05 under more optimistic assumptions ($\alpha_t = 0.27$ and $\alpha_c = 0.22$), and the red/top-right region shows combinations that yield bias greater than 0.05 even under optimistic assumptions.

Research Design:
Description of the research design.

(May not be applicable for Methods submissions)

Not applicable

Data Collection and Analysis:
Description of the methods for collecting and analyzing data.

(May not be applicable for Methods submissions)

Not applicable

Findings / Results:
Description of the main findings with specific details.

(May not be applicable for Methods submissions)

Not applicable

Conclusions:
Description of conclusions, recommendations, and limitations based on findings.

The model of attrition bias presented here, and the standard based on it, is an effort to protect consumers of research from drawing mistaken conclusions because of bias (in this case, attrition bias) in impact estimates. It is important to be clear about what this effort can and cannot do.

What we cannot do. We *cannot* claim that impact estimates from studies meeting the attrition standard are definitely free of attrition bias. Nor can we make precise probabilistic statements regarding the magnitude of attrition bias in studies that meet the attrition standard. This is primarily because (1) our empirical estimates of bias from past studies could be wrong (since we had to rely on pre-tests as a proxy for post-tests) and (2) our empirical estimates of bias from past studies may not apply to future studies (because, for example, future studies may involve different populations of students).

What we can do. We *can* claim that impact estimates from studies meeting the attrition standard are likely to be less influenced by attrition bias than estimates from studies that do not. Also, this approach allows us to systematically manage the tradeoff between overall and differential attrition, recognizing that studies with low differential attrition are preferable to studies with high differential attrition, holding the overall rate constant. Finally, this approach allows us to reduce the influence of bias in impact estimates in a way that is systematic and informed by data, which we believe is preferable to arbitrary cutoffs.

Appendices

Not included in page count.

Appendix A. References

References are to be in APA version 6 format.

Little, R.J.A., & D.B. Rubin (2002). *Statistical Analysis with Missing Data, 2nd Edition*. Hoboken, NJ: John W. Wiley and Sons.

Manski, C. F. (1990). "Nonparametric Bounds on Treatment Effects." *American Economic Review, Papers and Proceedings*. 80: 319-323.

Puma, Michael J., Robert B. Olsen, Stephen H. Bell, and Cristofer Price (2009). *What to Do When Data Are Missing in Group Randomized Controlled Trials (NCEE 2009-0049)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Rubin, D.B. (1976), "Inference and Missing Data." *Biometrika*, 63, 581-92.

Appendix B. Tables and Figures

Not included in page count.

FIGURE 1

TRADEOFFS BETWEEN OVERALL AND DIFFERENTIAL ATTRITION